

“A fehérjeszerkezetek eredete” c. posztdoktori OTKA projekt (PD 73096) szakmai záróbeszámolója

Bevezetés, előzmények, a kutatás célja

Az élőlényekben különböző fehérjék milliói látják el biokémiai, élettani feladataikat. E fehérjék szerkezetében azonban számos szabályosság ismerhető fel. Sokszor nagyon különböző funkciójú fehérjék egészen hasonló szerkezettel rendelkeznek. Ha a fehérjéket doménekre bontjuk szét (legfeljebb pár száz aminosav hosszúságú, globuláris egységek), s a domének szerkezetét megvizsgáljuk, azt látjuk, hogy legfeljebb néhány ezer, egymástól alapvetően különböző fehérjeszerkezet létezik. Ezeket az alapvető szerkezeteket nevezzük **foldoknak**. Egy fold lényegében nem más, mint a különböző másodlagos szerkezeti elemek (α -hélixek, β -szálak, hurkok) meghatározott térbeli elrendeződése.

A fehérjeevolúció vizsgálatai során hagyományosan egymáshoz felismerhető mértékben hasonló szekvenciákból indulnak ki. Az ilyen – egymással homológ – szekvenciák elemzése során filogenetikai törzsfák építhetők, így az egyes fehérjék evolúciós viszonyai föltérképezhetők. Az egymással homológ fehérjék szerkezete azonban mindig nagyon hasonló, doménszinten azonos folddal rendelkeznek. A fehérjeevolúció vizsgálatának hagyományos módszerei tehát nem adnak választ arra a kérdésre: vajon hogyan keletkeztek a különböző foldok¹⁻⁴? Milyenek a különböző foldok evolúciós viszonyai? Egymástól függetlenül keletkeztek-e, avagy származhat-e egyik fold a másiktól?

Fontos megfigyelés, hogy a fehérjék szerkezete jóval konzerválódottabb a szekvenciájuknál. Ha egy fehérjét kódoló gén megkettőződik, az utódgénnek két egyforma fehérjét fognak kódolni. Idővel azonban a véletlen mutációk hatására a két fehérje szekvenciája egyre különbözőbbé válik, s egy ponton már szekvenciaszinten felismerhetetlen lesz közös eredetük, míg szerkezetük alapvetően hasonló marad, foldjuk azonos marad vagy csak kisebb különbségek lesznek közöttük. A fehérjeszerkezet-jóslás egyik alapköve a *threading*, az a módszer, amellyel megkísérelhetjük felismerni két fehérje szerkezetbeli hasonlóságát (feltéve, hogy egyiknek a szerkezete, másiknak a szekvenciája ismert) akkor is, ha a szekvenciák már nagymértékben eltávolodtak egymástól, vagyis a „twilight zone”-ban vannak. Ugyanakkor hasonló fehérjeszerkezetek létrejöhetnek konvergens evolúció útján is, a szerkezeti hasonlóság tehát önmagában nem bizonyíték két fehérje evolúciós rokonságára. Többféle módszer ismeretes a homológok és az analógok megkülönböztetésére, ám ezeket teljes biztonsággal megkülönböztetni nem lehet.

Miután tehát a hagyományos filogenetikai analízis nem alkalmas a fehérjeszerkezetek evolúciós történetének feltárására, más módszerekre van szükség. Kutatási projektünk célja az volt, hogy ezen a téren mutasson fel előrelépést.

Természetesen a szakirodalomban már leírtak elképzeléseket, hipotéziseket a fehérjeszerkezetek eredetére és evolúciójára vonatkozóan. Az egyik ismert feltételezés szerint a mai, nagyobb fehérjeszerkezetek rövidebb, másodlagos vagy szupermásodlagos szerkezeti elemeknek megfelelő peptidek fúziója útján jöttek létre – ez az ADS (antecedent domain segments) hipotézis⁵. Egy másik elképzelés szerint létezhetnek olyan evolúciós átmenetek, amelyek a különböző foldokat egymásba képesek transzformálni⁶. Különböző szerkezetek kialakulhatnak egy kiinduló szerkezetből való divergencia útján², de nagyobb szerkezetek létrejöhetnek dinamikus, csak részben rendezett polipeptidekből is⁷.

A fehérjeszerkezetek evolúciójának kutatása szempontjából talán a leginformatívabbak az olyan fehérjeszerkezetek, amelyek ismétlődő szerkezeti elemekből állnak. Ezek voltaképpen

saját evolúciós történetükről árulkodnak. Egy feltételezés szerint az ilyen fehérjék egykori homooligomerek fúziójából jöhettek létre. Ennek szép példája a TIM-hordó, azaz az igen gyakori (ba)8 hordó evolúciója két félhordóból, melyre meggyőző bizonyítékok vannak⁸, illetve a β -háromlevél (trefoil) kifejlődése egy homotrimerből⁹.

Új foldokat generáló mechanizmus a **cirkuláris permutáció**¹⁰ is. Bár ennek során a foldot felépítő másodlagos szerkezeti elemek ugyanazok maradnak, a polipeptidláncon belüli sorrendjük megváltozik. A cirkuláris permutáció legvalószínűbb genetikai mechanizmusa a génduplikáció, fúzió, majd a terminálisok levágódása¹¹. A cirkuláris permutánsokat gyűjtő CPDB adatbázis jelenleg több mint 4000 nemredundáns, ismert szerkezetű cirkulárispermutáns-párt tartalmaz¹².

A fehérjék szerkezeti és funkcionális sokfélesége nagyot nőtt a többdoménos fehérjék megjelenésével¹³. Ezek az egyes domének duplikációjával, divergenciájával és rekombinációjával jönnek létre („domain shuffling” vagy fúzió). A különböző többdoménos elrendeződések sokféle geometriát¹⁴ és funkciót hoztak létre.

A negyedleges szerkezetek szintjén a háromdimenziós doméncsere az új szerkezetek kialakulásának egyik jelentős mechanizmusa^{15,16}. A doméncsere a homodimerek képződésének ideális módja, ugyanis nem kell új domén-domén interfészt kifejleszteni, ehelyett az interfész egyszerűen azáltal jön létre, hogy a két monomer egymás között kicseréli egy-egy, egymással ekvivalens részét, ezáltal a láncon belüli kontaktusokat velük egyenértékű, láncok közötti kontaktusokká alakítva¹⁷. Ezen az úton magasabbrendű oligomerek is létrejöhetnek, a monomerek egymás után kapcsolódásával¹⁸. Számos megfigyelés utal arra, hogy a doméncserének sok mai oligomer evolúciós történetében szerepe lehetett, még olyanokban is, amelyek nem látszanak doméncseréltnek¹⁹.

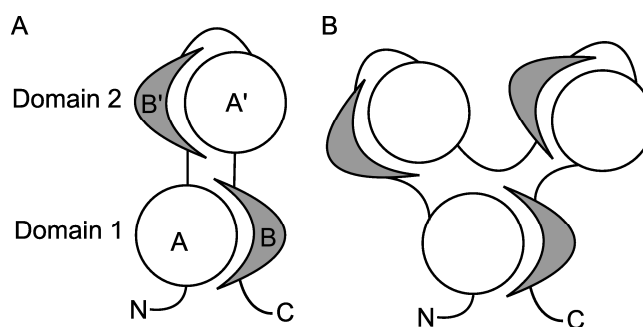
Kutatásunkban elsősorban az ismétlődő szerkezeti elemekből álló fehérjékre koncentráltunk, mivel ezek szerkezete árulkodik saját evolúciójukról. A projekt fő eredménye egy új fehérjecsoport felismerése, melyet **szegmenscserélt fehérjéknek** nevezünk el. A szegmenscserélt fehérjék evolúciós mechanizmusára vonatkozóan modelleket állítottunk fel, és nagyléptékű szerkezet- és szekvenciaanalíziseket végeztünk annak eldöntésére, melyik mechanizmusnak van domináns szerepe e fehérjék evolúciójában.

A kutatás melléktermékeként több olyan eredmény is született, mely a fő témához ugyan csak érintőlegesen kapcsolódik, ám értékes új felismeréshez vagy módszerhez vezetett el.

Szegmenscserélt fehérjék

Kutatásunk során az ismétlődő szerkezeti elemeket tartalmazó fehérjéket vizsgáltuk, s ennek során lettünk figyelmesek ezek egy olyan csoportjára, amelyet a szakirodalom a jelek szerint eddig még nem azonosított, nem vizsgált. Ezt a csoportot **szegmenscserélt fehérjéknek** neveztük el. Különösen ráirányította e fehérjékre figyelmünket az az esemény, hogy a fehérjeszerkezet-jósló módszereket tesztelő nemzetközi versenyen (CASP8) az egyik célfehérje (T0504) legjobb templátjaként a legtöbb foldfelismerő program egy szegmenscserélt fehérjét (2gf7A) ismert fel, holott – mint később kiderült – a célfehérje nem volt szegmenscserélt, így teljesen hibásan jósolt szerkezetek születtek.

Mik is a szegmenscserélt fehérjék? Olyan kétdoménos (esetleg kettőnél több doménos) fehérjék, amelyek „úgy néznek ki”, mintha egy doméncserélt homodimer (esetleg homooligomer) alegységeinek fúziójaként keletkeztek volna, a szekvenciák ezt követő divergenciájával. Tehát a fehérjerészek cseréje nem különálló alegységek között, hanem egyetlen láncon belüli domének között ment látszólag végbe. Az 1. ábra ezt szemlélteti:



1. ábra: Szegmenscserélt fehérjék sematikus ábrázolása. (A) kétdoménes (B) háromdoménes szegmenscserélt fehérje.

A leggyakoribb, kétdoménes szegmenscserélt fehérjék négy szegmensre oszthatóak: A, B', A', B. Az N- és C-terminálison lévő A és B szegmens alkot egy domént (Domain 1), és a hozzájuk szerkezetileg hasonló A' és B' szegmens egy másik, folytonos domént (Domain 2). A két doménben a szegmensek sorrendje (N-től C-terminálisig) fordított, tehát cirkuláris permutációi egymásnak. A lánc két fele szerkezetileg hasonló egymáshoz, ám mindkét doménből tartalmaznak részeket. Fontos különbség a doméncserélt homodimerek és a szegmenscserélt fehérjék között, hogy míg a doméncserélt homodimerek esetében az alegységeknek van nem doméncserélt (zárt) konformációja is (dinamikus egyensúly létezhet a doméncserélt és a nem doméncserélt forma között), addig a szegmenscserélt fehérjék esetében a két domén szekvenciája már erősen divergált, így a szegmensek nem tudnak „visszacserélődni”. Az A szegmens pl. csak a B-vel kompatibilis, a B'-vel már nem.

Szegmenscserélt fehérjék azonosítása

Ahhoz, hogy a szegmenscserélt fehérjéket tanulmányozzuk, először is össze kell gyűjteni őket. Ezért kifejlesztettünk egy algoritmust, melynek révén ezek a háromdimenziós szerkezet alapján azonosíthatók. Az eljárást a Protein Data Bank (PDB) egy általunk összeállított részhalmazán futtattuk; a részhalmazt úgy állítottuk össze, hogy szerkezetileg reprezentálja a teljes PDB-t. Így a teljes PDB kb. 150 000 lánc helyett csak 21 938 láncra kellett futtatni az algoritmust.

Az algoritmus a TM-align nevű programon alapul²⁰, mely hasonló szerkezeteket illeszt egymásra, és kiszámítja a szerkezetek hasonlóságát jellemző, ún. TM-score-t. Ennek értéke 0 és 1 között van, és igazolt, hogy a 0,5-nél nagyobb TM-score azonos vagy nagyon hasonló foldot jelent.

Az eljárás során a szerkezeteket doménekre bontjuk a Domainparser2 programmal²¹, majd megkeressük azokat a fehérjéket, amelyek pontosan 2 doménből állnak és az egyik domén folytonos, a másik pedig nem folytonos. Ezután a nem folytonos domén két darabját felcseréljük és TM-align segítségével a folytonos doménre illesztjük. Amennyiben a TM-score nagyobb 0,5-nél és mindkét szegmens legalább 2/3-os lefedettséggel illeszkedik, akkor szegmenscserélt fehérjéről van szó. Többdoménes szegmenscserélt fehérjék keresésénél hasonlóan járunk el, de mindegyik folytonos domén esetében megköveteljük a leírt feltételeknek megfelelő illeszkedést. Az illesztés alapján a folytonos domént is felosztottuk, így minden fehérje esetében definiáltuk az 1. ábra szerinti szegmenseket.

A módszer segítségével 32 szegmenscserélt fehérjét azonosítottunk, melyek szerkezeti hasonlóság alapján 18 fehérjecsaládba sorolhatóak. Ezeket az 1. táblázat sorolja fel, a 2. ábra pedig bemutat közülük néhány jellemző szerkezetet.

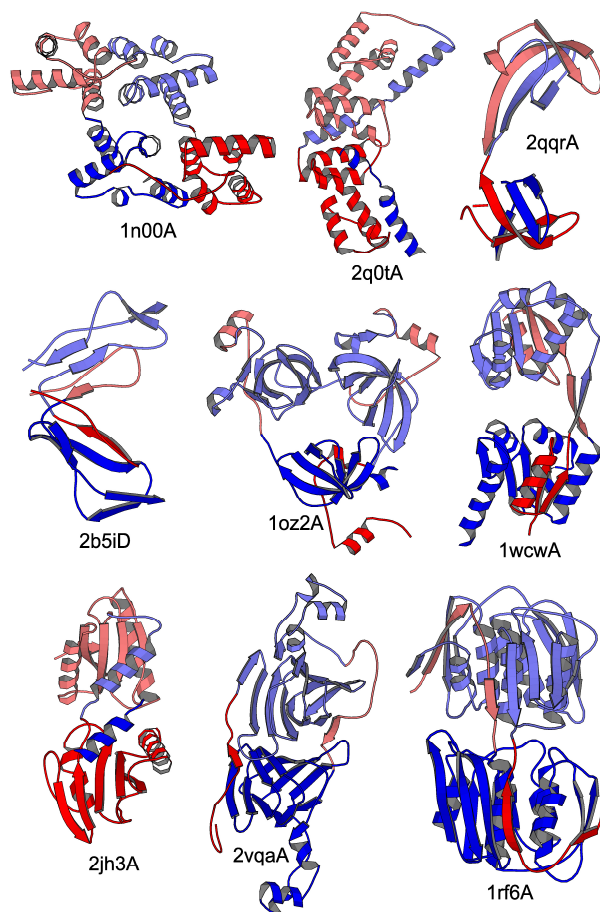
Egy másik, megengedőbb definíciót is használtunk, melynek során a szegmensenkénti 2/3-os lefedettséget nem követeltük meg. Ez 12 további fehérjecsaládot eredményezett, vizuálisan ellenőrizve azonban ezek nem bizonyultak meggyőző módon szegmenscserélt

fehérjéknek, noha lehetnek szegmenscserélt fehérjék leszármazottai. Mindazonáltal ennek alapján úgy becsüljük, hogy a többdoménos fehérjék kb. 12%-ának evolúciós történetében szerepet játszhat a szegmenscsere.

1. táblázat: Szegmenscserélt fehérjék és tulajdonságaik

Domén, fold, család neve	Funkció	PDB láncok	Átlagos doménméret / Csere mértéke
Főleg-α foldok			
Orthogonal bundle, annexin fold	phospholipid binding	1n00A 1dk5A	158 / 53%
Up-down bundle, AhpD-like fold	lyase, decarboxylase	2q0tA	127 / 70%
Főleg-β folds			
SH3-like, two MBT repeats	transcriptional regulator	2r58A 2bivA 1oz2A ^c	106 / 23%
SH3-like, Jumonji domains	oxidoreductase, demethylase	2qqrA	58 / 52%
2-layer β -sandwich variant ^d	hydrolase, galactosidase	3d3aA	127 / 14%
Complement control module	receptor	2b5iD	61 / 30%
Double-stranded β -helix; RmlC-like cupins	oxidoreductase	1y3tA	165 / 12%
3-rétegű $\alpha\beta\alpha$ szendvicsek			
Periplasmic binding protein-like II	transcriptional regulator	2ql3A 3hhfA 1ixcA	100 / 75%
HemD-like	lyase	1wcwA 1jr2A	128 / 27%
PrpR receptor domain-like	transcriptional regulator	2q5cA	93 / 80%
NAD(P)-binding Rossmann fold	oxidoreductase, dehydrogenase	2et6A	291 / 80%
Rossmann fold variant 1 ^d	transferase	2hcrA	152 / 87%
Rossmann fold variant 2 ^d	unknown	2jh3A	132 / 82%
Egyéb $\alpha\beta$ fehérjék			
$\alpha\beta$ -prism (6 repeats of IF3 fold)	transferase	1rf6A 2yvwA 1g6sA 1ejdA 2pqcA 2o0bA 2rl1A	212 / 9%
CBS domain pair ($\alpha\beta\beta\alpha$ sandwich)	adenosyl binding	1yavA 3hf7A 2emqA	66 / 22%
Double-stranded β -helix + α -helices ^d	metal binding	2vqaA	178 / 13%
$\alpha\beta$ -roll, diaminopimelate epimerase-like	unknown (isomerise?)	2h9fA	190 / 89%
WWE domain	signaling	2a90A	77 / 91%

A fehérjék között mindössze egyet találtunk, amelynek 2-nél több doménje van: a háromdoménos 1oz2A. A talált foldok mindegyik fehérjeosztályt lefedik, különösen a 3-rétegű $\alpha\beta\alpha$ szendvicsek fordulnak elő nagy számban. Funkcionálisan is nagy a sokféleség: enzimek, transzkripciós szabályozófehérjék, jelátviteli és más kötőfehérjék is előfordulnak. Méretük 58-291 aminosav, a szegmenscsere mértéke 9-91%. A két domén közötti szekvenciaazonosság átlagosan 19%, mindössze 4 esetben haladja meg a 30%-ot. Itt jelezzük, hogy a 12% feletti szekvenciaazonosság ebben az esetben már homológiára utal, tehát a két domén evolúciós rokonságához kevés kétség férhet.



2. ábra: Néhány szegmenscserélt fehérje szerkezete. Az egymásnak megfelelő szegmensek hasonló színűek.

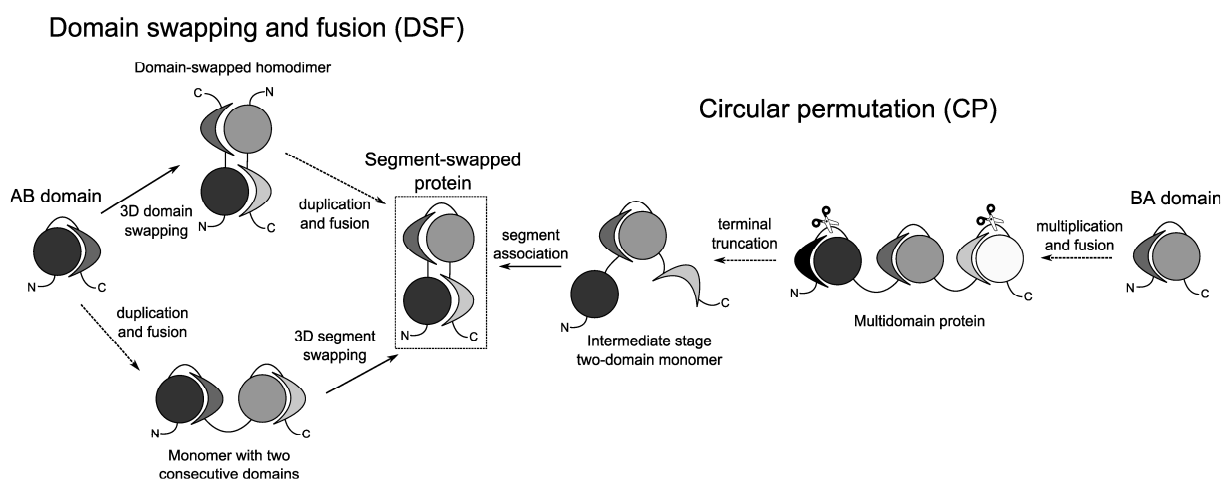
Evolúciós mechanizmusok

Kutatásunk következő kérdése: **milyen úton alakulhattak ki a szegmenscserélt szerkezetek?** A szerkezet alapján természetesen adja magát az a feltételezés, hogy ilyen fehérjék egy doméncserélt homodimer alegységeinek (génduplikációt követő) fúziójával, majd ezt követő szekvenciadivergenciával jöhetnek létre. Kétségtelenül ez a legegyszerűbb mechanizmus, azonban találtunk egy másik, szintén eléggé egyszerű mechanizmust is. A két legegyszerűbb mechanizmus tehát (3. ábra):

1. Doméncsere és fúzió (DSF) mechanizmus: a kiinduló domén képes doméncserélt homodimereket képezni. Ennek génje duplikálódik és a két gén fuzionál, így egy láncon belül jön létre a doméncsere (amit így már szegmenscserének nevezünk). A

szekvencia két fele ezután divergálódik, a szegmenscserélt konformáció ily módon genetikailag stabilizálódik.

2. Cirkuláris permutáció (CP) mechanizmus: A kiinduló domén legalább triplikálódik és fuzionál, tehát létrejön egy legalább három kópiát tartalmazó lánca. Ennek végei levágódnak és a maradék részek egy új, nem folytonos domént alkotnak.



3. ábra: A szegmenscserélt fehérjék keletkezésének két legegyszerűbb mechanizmusa

A CP mechanizmus alapötletét a β -propeller fold adta, mely szintén szegmenscseréltnek tekinthető, ha egy lapátot egy doménnek tekintünk. Ennek különböző változatai egymás cirkuláris permutációi. Természetesen más evolúciós mechanizmusok is elképzelhetők,

azonban ezek a legtakarékosabbak abban az értelemben, hogy a legkevesebb genetikai lépést igénylik, és ezért a legvalószínűbbek.

Felmerül ezek után a kérdés: ha ezt a két mechanizmust fogadjuk el mint legvalószínűbbeket, melyik fordul elő gyakrabban a valóságban? Ennek felderítésére háromféle tesztet végeztünk el.

Az első teszt alapvető észrevétele az, hogy – mint a 3. ábra mutatja – ha a jelenleg megfigyelt szegmenscserélt fehérje szerkezetét az ABAB séma írja le, akkor azt ehhez elvezető DSF mechanizmusnak egy AB, míg a CP mechanizmusnak egy BA szerkezetű doménből kellett kiindulnia. Az a kérdés, hogy melyik mechanizmus hozta létre a fehérjét, tehát átfogalmazható így: az AB vagy a BA típusú domén ősi-e? Végül soron tehát a kétféle domén relatív korát kell meghatároznunk. A szakirodalom ismer néhány módszert domének relatív korának meghatározására^{1,22}, a legegyszerűbb ilyen módszer szerint egy domén annál idősebb, minél több élőlény genomjában fordul elő. Ennek nyomán megszámláltuk mind a 18 szegmenscserélt fehérjecsalád esetében az AB, ill. BA fold előfordulását mint a PDB-ben (szerkezeti összehasonlítás útján), mind 22 genomban (rejtett Markov-modellekkel²³ történő szekvenciaösszehasonlítás útján). Az eredményeket a 2. táblázat mutatja.

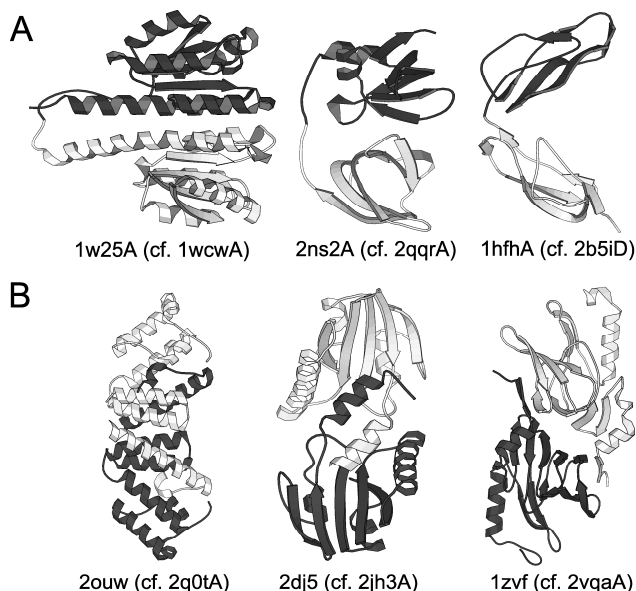
2. táblázat: Szegmenscserélt fehérjék két doménjéhez tartozó analógok előfordulása a PDB-ben (forrásorganizmusok száma szerint), ill. 22 genomban.

Csoport	Központi fehérje	A ReprPDB-ben AB-típus / BA-típus	22 genomban AB-típus / BA-típus
Főleg-α	1n00A	0 / 0	0 / 1
	2q0tA	1 / 0	11 / 1
Főleg-β	2r58A	0 / 1	2 / 0
	2qqrA	43 / 0	6 / 0
	3d3aA	37 / 11	0 / 0
	2b5iD	4 / 0	4 / 3
	1y3tA	2 / 0	5 / 0
3-rétegű $\alpha\beta\alpha$ szendvics	2ql3A	13 / 1	9 / 4
	1wcwA	137 / 98	21 / 10
	2q5cA	43 / 24	20 / 8
	2et6A	0 / 0	6 / 1
	2hcrA	105 / 66	9 / 5
	2jh3A	67 / 10	4 / 10
Egyéb $\alpha\beta$	1rf6A	0 / 0	0 / 0
	1yavA	0 / 1	20 / 1
	2vqaA	0 / 0	2 / 6
	2h9fA	0 / 0	0 / 0
	2a90A	0 / 0	2 / 1

A legtöbb esetben az AB fold előfordulása volt gyakoribb mindkét teszt szerint, ami a DSF mechanizmus dominanciáját támasztja alá. A CP mechanizmus a háromdoménos 1oz2A esetében valószínűsíthető.

Találtunk néhány olyan fehérjét is, amely két, egymás után AB domént tartalmazott – ez tulajdonképpen a szegmenscserélt fehérje „cseréletlen” verziójának tekinthető (4A. ábra), szintén a DSF mechanizmust alátámasztva.

Egy másik teszt során olyan doméncserélt homodimereket kerestünk, amelyek megfelelnek valamely létező szegmenscserélt fehérjének, csak épp a két domént két alegység „helyettesíti”. Öt szegmenscserélt fehérjére találtunk ilyen analógot (ezekből hármat a 4B ábra mutat be). Ezek létezése a DSF mechanizmust támasztja alá.



4. ábra (A) Két AB domént tartalmazó fehérjék (B) Homodimerikus analógok.

A harmadik teszt során abból az észrevételből indultunk ki, hogy amennyiben a két mechanizmus különböző változatokat hoz létre attól függően, hogy hogyan oszlik fel két szegmensre a kiinduló domén, akkor a DSF mechanizmus által létrehozott változatokban az 1. domén állandó marad és a 2. domén változik, míg a CP által létrehozottakban éppen fordítva. Ennek nyomán tehát a szegmenscserélt fehérjék különböző változatait kerestük, és meg is állapítottuk, hogy a 18 azonosított család közül kettő-kettő egymás változatainak tekinthető (1wcwA:2hcrA és 2r58A:2qqrA). Mindkét esetben az 1. domén hasonló, a 2. különböző szerkezetű, ami ismét a DSF mechanizmus működését támasztja alá.

Összefoglalva tehát azt találtuk, hogy a DSF a domináns evolúciós mechanizmus. Ez összhangban van azzal, hogy a DSF kevesebb genetikai lépést igényel, mint a CP.

Szegmenscsere és funkció

A szegmenscserélt fehérjéket funkció szempontjából is elemeztük, és azt találtuk, hogy ezek két csoportra oszthatók:

1. A két domén egymástól függetlenül, hasonló funkciót tölt be, ligandumok mindkét doménhez, külön kötődnek.
2. A ligandumok a két domén közötti hasadékba kötődnek, a funkcióhoz lényeges a két domén relatív csuklómozgása.

A szegmenscsere következménye, hogy a két domént nem egy, hanem két összekötő, linker régió kapcsolja össze. Ennek nyomán valószínűsítjük, hogy ennek funkcionális jelentősége lehet, ti. a két linker jelenléte hatékonyabb, egyetlen tengely körüli elfordulást tesz lehetővé, mintha csupán egy linker lenne, amely körül szabadon foroghatnak a domének.

Az eredmények jelentősége, kísérleti ellenőrzés, publikáció

A szegmenscserélt fehérjék azonosításával és elemzésével egy új, eddig fel nem ismert, saját jogán nem vizsgált fehérjecsoport létezésére hívtuk fel a figyelmet. Ezáltal számos további kutatás előtt nyitottuk meg az utat.

Az eredmények felhasználásával kísérletek is tervezhetőek, melyekkel eldönthető, hogy a leírt mechanizmusok a valóságban is működőképesek-e. Megterveztünk egy kísérletsorozatot a szegmenscserélt-konzekutív konformációk közötti átkapcsolás lehetőségességének ellenőrzésére egy CCP doménpárt tartalmazó fehérjében. A kísérletek kivitelezése jelenleg folyik.

A szegmenscserélt fehérjékkel kapcsolatos kutatásunk eredményeit publikáció formájában írásba foglaltuk. Ez jelenleg beküldött állapotban, bíráló alatt van egy vezető molekuláris biológiai folyóiratban (Szilágyi et al.: Intra-chain segment swapping spawns the evolution of new multidomain architectures. *Submitted.*).

Egyéb eredmények

Doménzáródás IPMDH-ban

Az izopropil-malát dehidrogenáz (IPMDH), mely intézetünkben a hagyományos enzimológiai kutatások egyik tárgya, valójában szegmenscserélt monomerekből álló dimer, noha szegmenscserélt volta nem nyilvánvaló az N-terminális „ragasztott” két, hosszú hélix miatt. A szegmenscsere eredménye a két linker régió a domének között. Kísérleti úton vizsgáltuk a domének csuklószerű mozgását, ill. ennek szerepét az enzim katalitikus tulajdonságainak hőmérsékletfüggésében. A munka eredményét a *Biophysical Journal*-ban publikáltuk (Hajdú et al, *Biophys J*, 96:5003).

Átmenet a rendezettség és a rendezetlenség között

A fehérjeevolúció egyik fő elmélete szerint a nagyobb szerkezetek valószínűleg rövidebb peptidekből álltak össze. A rövidebb peptidek ugyanakkor nem feltétlenül rendezett szerkezetek, valószínűleg többnyire inkább dinamikusan változó szerkezetű láncok. Ezért kérdéses, hogyan képezhetnek rendezett szerkezetű konglomerátumot. Ezt a kérdést vizsgáltuk HP (hidrofób-poláros) modellek segítségével. A tanulmány egyik fő konklúziója, hogy a fehérjeméretnek meghatározó szerepe van a rendezettség létrehozásában: adott aminosav-összetétel mellett a rövidebb láncok lehetnek rendezetlenek, a hosszabbak pedig már rendezettek. A pusztán méretnövekedésnek tehát rendező hatása van. Ezt a munkát a *Biophysical Journal*-ban publikáltuk (Szilágyi et al.: The twilight zone between protein order and disorder. *Biophys J*, 95, 1612). A HP modellekkel történő további munkánk a rendezetlen-rendezett átmenet részletesebb jellemzését szolgálja, ebből publikáció előkészületben van.

DNS-kötő fehérjék azonosítása

Az izraeli Nir-Ben Tal csoportjával együttműködésben egy DNS-kötő fehérjék azonosítására szolgáló módszert fejlesztettünk ki, mely többek között evolúciósan konzerválódott felszíni „foltok” azonosításán alapul. Az eredményeket publikáltuk (Nimrod et al, *J Mol Biol* 387:1040; Nimrod et al, *Bioinformatics* 26:692).

Galektinek egy csoportjának evolúciója

Thán Gáborral (Detroit, USA) együttműködésben a galektinek egy alcsaládjának evolúcióját vizsgáltuk bioinformatikai eszközökkel. Az eredményeket publikáltuk (Than et al, *PNAS* 106:9731).

Fehérjeszerkezet-predikció

Yang Zhanggal (Ann Arbor, USA) együttműködésben a fehérjeszerkezet-predikciót segítő eljárásokat fejlesztettünk ki. Ezek egyike az az eljárás, amely a szekvencia alapján, evolúciós információt is felhasználva jósol meg láncon belüli kontaktusokat. Az eredményeket publikáltuk (Wu et al, *Structure*, in press).

Entrópiaszámítás

Kutatásaink során szimulációkat végeztünk egyszerűsített fehérjemodellekkel (ún. diszkrét molekuladinamika) a rendezetlenség és rendezettség közötti átmenetek leírására. Ehhez szükségünk van az entrópiaváltozások megbízható számítására is. Az irodalomban fellelhető módszereket áttekintve kifejlesztettünk egy új módszert, amely a szimuláció során nyert trajektóriák valószínűsűrsűrűség-függvények Gauss-keverékekkel történő modellezésén alapul. A publikációt beküldtük, bírálattal áll (Gyimesi et al, *submitted*).

Irodalomjegyzék

1. Abeln, S.; Deane, C. M. *Proteins* 2005, 60(4), 690.
2. Deeds, E. J.; Shakhnovich, E. I. *Adv Enzymol Relat Areas Mol Biol* 2007, 75, 133.
3. Koonin, E. V. *Biol Direct* 2007, 2, 21.
4. Rost, B. *Curr Opin Struct Biol* 2002, 12(3), 409.
5. Lupas, A. N.; Ponting, C. P.; Russell, R. B. *J Struct Biol* 2001, 134(2-3), 191.
6. Taylor, W. R. *Curr Opin Struct Biol* 2007, 17(3), 354.
7. Tokuriki, N.; Tawfik, D. S. *Science* (80-) 2009, 324(5924), 203.
8. Lang, D.; Thoma, R.; Henn-Sax, M.; Sterner, R.; Wilmanns, M. *Science* (80-) 2000, 289(5484), 1546.
9. Ponting, C. P.; Russell, R. B. *J Mol Biol* 2000, 302(5), 1041.
10. Lindqvist, Y.; Schneider, G. *Curr Opin Struct Biol* 1997, 7(3), 422.
11. Vogel, C.; Morea, V. *Bioessays* 2006, 28(10), 973.
12. Lo, W.-C.; Lee, C.-C.; Lee, C.-Y.; Lyu, P.-C. *Nucleic Acids Res* 2009, 37(Database issue), D328.
13. Vogel, C.; Bashton, M.; Kerrison, N. D.; Chothia, C.; Teichmann, S. A. *Curr Opin Struct Biol* 2004, 14(2), 208.
14. Bashton, M.; Chothia, C. *J Mol Biol* 2002, 315(4), 927.
15. Bennett, M. J.; Choe, S.; Eisenberg, D. *Proc Natl Acad Sci U S A* 1994, 91(8), 3127.
16. Bennett, M. J.; Schlunegger, M. P.; Eisenberg, D. *Protein Sci* 1995, 4(12), 2455.
17. Gronenborn, A. M. *Curr Opin Struct Biol* 2009, 19(1), 39.
18. Jaskólski, M. *Acta Biochim Pol* 2001, 48(4), 807.
19. Bennett, M. J.; Eisenberg, D. *Structure* 2004, 12(8), 1339.
20. Zhang, Y.; Skolnick, J. *Nucleic Acids Res* 2005, 33(7), 2302.
21. Guo, J.-t.; Xu, D.; Kim, D.; Xu, Y. *Nucleic Acids Res* 2003, 31(3), 944.
22. Winstanley, H. F.; Abeln, S.; Deane, C. M. *Bioinformatics* 2005, 21 Suppl 1, i449.
23. Söding, J. *Bioinformatics* 2005, 21(7), 951.